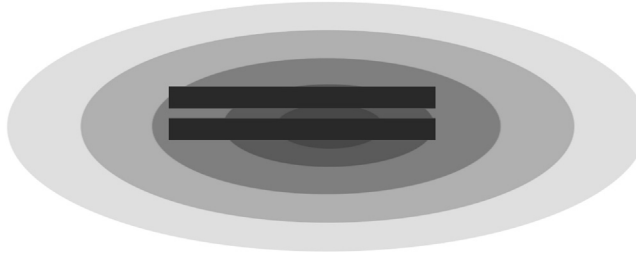
CHAPTER 7

Solenoid — An amyloid under control

Mateusz Banach[1], **Irena Roterman**[1]
[1]Department of Bioinformatics and Telemedicine, Jagiellonian University—Medical College, Krakow, Poland

Contents



The diagram illustrates a structure which is consistent with the fuzzy oil drop model despite containing a solenoid fragment, characterized by linear propagation of alternating bands of variable hydrophobicity. This unusual structure, when "packaged" in a soluble envelope, is protected against unrestricted aggregation.

The solenoid is a supersecondary structure shaped like a tubule, with three or four "walls" consisting of short β-strands generating elongated parallel one to each other β-sheets. This unusual structure is sometimes also called a β-helix. The symbol "β" refers to the presence of β-strands in each segment, while "helix" describes the general characteristics of the

From Globular Proteins to Amyloids
ISBN: 978-0-08-102981-7
https://doi.org/10.1016/B978-0-08-102981-7.00010-5

solenoid as a whole, since solenoids are usually slightly twisted about their main axis.

An α-helix is stabilized by a network of hydrogen bonds which run parallel to its axis. These bonds join every third or fourth amino acid. In a solenoid, numerous hydrogen bonds are found within the system of β-strands forming β-sheets which make up the walls of the structure. Each of these bonds links residues which are much further apart in the chain than in the case of an α-helix (depending on the overall breadth of the solenoid and on the optional presence of exposed loops between each unit segment).

Entering "solenoid" in the PDB search interface produces two groups of proteins: antifreeze proteins and lyases. In the former group, solenoid fragments tend to be regularly structured, while in lyases they are often accompanied by numerous additional elements such as helical folds, loops (often quite long), random coils and even domain-like fragments. Thus, we will illustrate our discussion of solenoids with two sample proteins representing both groups.

The structural peculiarity of solenoids immediately raises the question — how can such a structure emerge in the first place? It turns out that each solenoid contains bands of bands of variable hydrophobicity which propagate along its main axis. The arrangement of bands appears to promote formation of long tubular structures, consisting of three or four β-sheets (four or sometimes three) with a more or less symmetrical cross section.

In a helix the radius of curvature is far smaller than in β-solenoids. Additionally, helices are more regular than solenoid fragments. Based on specific examples we will attempt to show that alignment of bands characterized by variable hydrophobicity is the main driving force behind formation of elongated structural forms. The linear propagation of bands is characterized by the possible no-limit propagation in contrast to globular 3 D Gauss-like structural forms. This is why solenoid as any other structure based on linear propagation of any character requires specific "caps" to stop unlimited propagation. This aspect will be discussed using the examples of proteins with solenoids in their structures.

Antifreeze proteins represented by protein extracted from *Flavobacterum frigoris PS1* found in the Antarctic region

As an example of the antifreeze family, we will consider a protein extracted from *Flavobacterum frigoris PS1* found in the Antarctic region

(PDB ID: 4NU2) [1]. This protein was selected due to its strong topological resemblance to the lyase which will be presented later on.

The antifreeze protein (4NU2) consists of a solenoid fragment along with several outlying fragments forming a package: a short "stop" helix (also referred to as "cap") attached to a terminal section of the solenoid and preventing unrestricted complexation, another helix which runs parallel to the solenoid. The protein consists also several random coil fragments.

The molecule as a whole does not conform to the Gaussian distribution of hydrophobicity (Table 7.1). Likewise, the solenoid itself (which in this case consists of three β-sheets) is strongly discordant. The sheet identified as

Table 7.1 FOD parameters (RD values and correlation coefficients) for antifreeze protein (4NU2). Results are listed for the complete molecule, for the β-sheets comprising the solenoid, for the solenoid as a whole and for other fragments of the chain, including selected secondary folds. Non-solenoid − parts of the chain which do not comprise the solenoid; 70% of non-solenoid − part of the chain which remain accordant with the expected distribution of hydrophobicity (detailed description given in the text). Values listed in boldface indicate discordance.

Antifreeze protein (4NU2)	Fragment	RD		Correlation coefficient		
		T-O-R	T−H	HvT	TvO	HvO
Chain	61−276	0.669	0.517	0.290	0.470	0.642
Solenoid						
β-sheet	**73−77***	**0.744**	**0.518**	**0.113**	**0.402**	**0.483**
β-sheet	81−84*	0.338	0.392	0.538	0.802	0.816
β-sheet	88−90*	0.718	0.550	0.181	0.169	0.805
β-sheet	166−167*	0.725	0.667	0.829	0.457	0.623
β-sheet	**88−90 + 166−167**	**0.808**	**0.694**	**0.163**	**0.157**	**0.745**
Complete		**0.660**	0.437	0.302	0.490	0.675
Other fragments						
Stop − random coil	104−112	0.298	0.110	0.426	0.761	0.828
Stop − helix	153−164	0.337	0.197	0.478	0.741	0.271
Non-solenoid		**0.616**	0.432	0.246	0.370	0.644
70% non-solenoid		0.470	0.320	0.352	0.632	0.658
Parallel helix	128−148	0.479	0.477	0.452	0.632	0.697
RC (parallel)	61−73	0.292	0.375	0.389	0.810	0.601
SS-bond	107−124	**0.672**	0.494	0.233	0.466	0.655

* only numbers of residues belonging starting (first) Beta-strand in the Beta sheet are given

73—77 (values corresponding to the location of its initial β-strand) is particularly amyloid-like, with high RD and dominant HvO correlation coefficient.

The solenoid is defined in this work as set of β-strands forming the helical form. The β-strands together with one additional residue at each end of strand is taken to define the solenoid part of the protein. This notation will be used all over this work.

A single disulfide bond (107C—124C) is present and appears to structurally reinforce the protein.

As already discussed, the solenoid is a linear structure. Producing such a structure calls for a mechanism which promotes elongation of a tubular form. According to observations involving other antifreeze proteins which include solenoid fragments, we assume that this mechanism is provided by a specific distribution of hydrophobicity, deviating from the monocentric Gaussian in favor of a linear pattern where bands of high and low hydrophobicity are arranged in an alternating fashion.

The profiles shown in Fig. 7.1A make it easy to identify the solenoid, which is characterized by a sinusoidal pattern of hydrophobicity, along with strong deviations from T. The C-terminal fragment, where

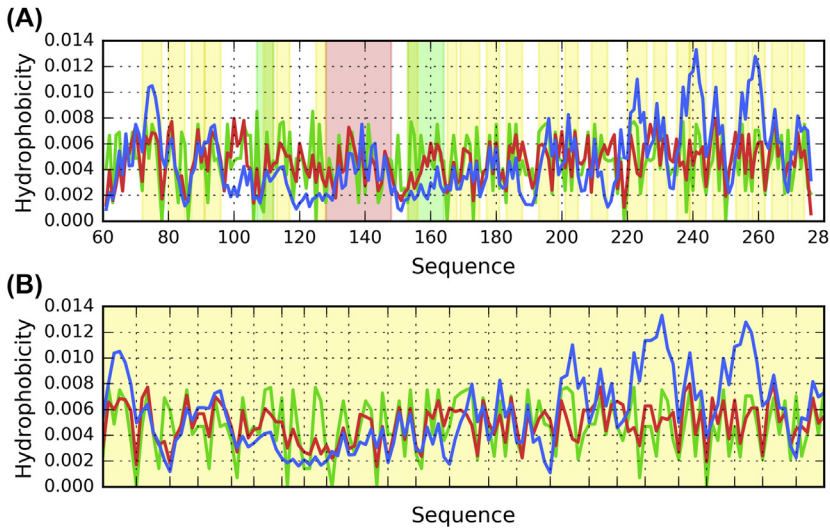


Fig. 7.1 Theoretical (T, blue), observed (O, red) and intrinsic (H, green) hydrophobicity distribution profiles for antifreeze protein (4NU2). (A) Complete molecule, (B) Solenoid fragments (β-strands given in Table 7 1, extended by 1 residue in each direction). Background colors highlight the position of solenoid (yellow), parallel helix (red) and "stop" fragments (green). To avoid displaying multiple gaps between the fragments of the solenoid, profiles in B are conflated into contiguous sequence. However ticks on the horizontal axis can be used to determine the location of those fragments within the complete sequence.

hydrophobicity is expected to remain strong, does not differ significantly from the remainder of the chain and is therefore discordant. Plotting the distribution of hydrophobicity for the solenoid itself (Fig. 7.1B) reveals a characteristic sinusoidal pattern composed of alternating bands. Clearly, this pattern does not correspond to the theoretical monocentric distribution of hydrophobicity. Another characteristic property of solenoids is the strong convergence between O and H (Fig. 7.1B, further confirmed by the data in Table 7.1). This strong correlation HvO with low correlation for HvT and especially TvO suggests the dominating role of intrinsic hydrophobicity of each residue. It is in contrast to expected uni-centric hydrophobic core formation which requires the synergy of all residues.

Taken together, it will shown later on that the structural properties of solenoid fragments resemble those which characterize amyloid fibrils.

This linear propagation — especially in the context of β-sheets — is capable of unrestricted growth, given that both edges of the solenoid fragment are able to form H-bond systems. Consequently, a dedicated "stop" signal is required to prevent the solenoid from growing without bound. Following analysis of a large set of solenoid-containing proteins we find that the signal assumes the form of a short helical "cap", which is typically highly consistent with the fuzzy oil drop model (low RD values). This means that the cap must be amphipathic: on its underside it remains compatible with the outermost segment of the solenoid, while on the outside it ensures suitable conditions for interaction with the aqueous solvent. Once in place, it effectively arrests further linear propagation. In later sections we will again refer to capping helices as a potential targets in the search for new drugs capable of preventing growth of amyloid structures. It will be shown in the Chapter 10.

Another characteristic property of solenoid-containing proteins is the presence of a long helix which runs parallel to the solenoid fragment and remains consistent with the Gaussian distribution. This helix also mediates entropically advantageous contact with the aqueous environment: it promotes solubility by preventing face-on complexation (which might otherwise be possible given the solenoid's discordant status). The helix is typically closely aligned with a discordant β-sheet — in our case with the sheet identified as 72—77.

Other components of the "packaging" (which refers to fragments other than the solenoid itself) exhibit variable FOD status. Altogether, 70% of the non-solenoid part of the chain remains consistent with the model and appear to promote solubility. We identify this part by progressively eliminating

discordant residues until we obtain RD < 0.5. Some additional discordances may be explained by the dynamic stereochemistry of outlying loops, which may locally deviate from the Gaussian distribution. According to the fuzzy oil drop model, antifreeze proteins perform their function by enforcing a nonstandard structural arrangement of nearby water molecules, disfavoring the formation of ice crystals. This disruption – while not yet precisely defined – may explain the need for a solenoid fragment, with its peculiar distribution of hydrophobicity. Experimental research suggests that water effectively "levitates" on top of hydrophobic surfaces [2], while the structure of the solvent adjacent to hydrophilic bands is likely determined by interactions between H_2O molecules and polar groups. Interestingly, the mobility of water molecules is observed to increase in closeness, proximity to antifreeze proteins [3]. A broader selection of antifreeze proteins is analyzed in Ref. [4].

An alternating pattern of variably hydrophobic bands appears optimal given the expected role of antifreeze proteins. In the Chapter 5. Which discuss the properties of FOD-compliant molecules we mention, among others, type III antifreeze proteins. Their surface is entirely composed of polar groups, which also affects the surrounding solvent and prevents crystallization. A similar mechanism may be proposed for other small organic molecules such as saccharides, lipids or even salt, which is a popular de-icing agent. Fig. 7.2 illustrates the location of various fragments listed in Table 7.1.

Lyase family represented by pectate lyase

The lyase family is represented by pectin and pectate lyase E.C. 4.2.2.2 isolated from *Bacillus subtilis* [5]. This enzyme catalyzes eliminative cleavage

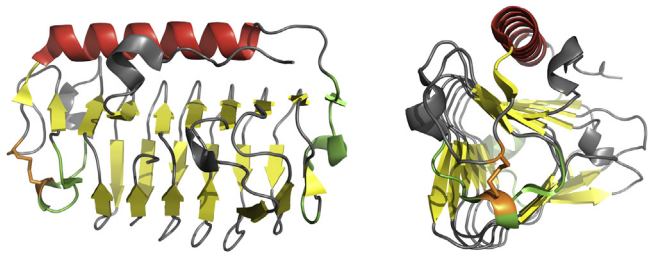


Fig. 7.2 3D presentation of antifreeze protein (4NU2) in two perspectives. Colors distinguish different fragments of the molecule: solenoid – yellow, parallel helix – red, "stop" fragments – green, disulfide – orange.

of pectate to give oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at their non-reducing ends. The structure is listed in PDB under ID 1BN8.

The presence of a Ca^{2+} ion close to an arginine residue appears important for catalysis since it is conserved across all pectin and pectate lyases. The protein consists of 399 residues and contains no disulfide bonds. Catalytic residues are identified as 184D, 206R and 279R.

Lyase (1BN8) is a very complex enzyme, both from the point of view of topology and intended function. Consequently, our analysis must address several distinct issues: its enzymatic activity and specific conformational properties, which are characterized in Table 7.2.

The molecule as a whole appears to lack a central hydrophobic core. Comparison of correlation coefficients indicates strong involvement of intrinsic hydrophobicity in shaping the tertiary conformation of lyase (1BN8) (Table 7.2 and Fig. 7.3).

Eliminating catalytic residues produces a slight decrease in RD, suggesting that these residues deviate from the theoretical distribution (although not very strongly). Further analysis reveals that discordances are centered

Table 7.2 FOD parameters (RD values and correlation coefficients) for lyase (1BN8). Results are listed for the complete molecule, for the β-sheets comprising the solenoid, for the solenoid as a whole and for other fragments of the chain, including selected secondary folds. Values listed in boldface indicate discordance. Amyloid-like fragments: bold − strong discordance between T and O including these with biased relation of correlation coefficients − the relation HvO dominating.

		RD		Correlation coefficient		
Lyase (1BN8)	**Fragment**	**T-O-R**	**T-O-H**	**HvT**	**TvO**	**HvO**
Chain	1−399	**0.684**	**0.559**	0.194	0.346	0.749
No E res.	184D, 206R, 279R	**0.680**	**0.553**	0.211	0.361	0.749
184D	±5 aa	**0.529**	**0.553**	−0.064	0.333	0.775
206R	±5 aa	0.482	**0.654**	0.360	0.711	0.880
279R	±5 aa	**0.505**	0.353	0.212	0.231	0.865
All cat. res.	**162−280**	**0.563**	0.425	0.188	0.418	0.730
Solenoid						
β-sheet	32−35*	**0.696**	0.361	0.350	0.414	0.709
β-sheet	61−63*	**0.705**	**0.538**	0.132	0.079	0.892
β-sheet	**122−125***	**0.730**	**0.628**	**−0.135**	**−0.177**	**0.655**
Complete		**0.727**	**0.529**	**0.121**	**0.148**	**0.773**

(*Continued*)

Table 7.2 FOD parameters (RD values and correlation coefficients) for lyase (1BN8). Results are listed for the complete molecule, for the β-sheets comprising the solenoid, for the solenoid as a whole and for other fragments of the chain, including selected secondary folds. Values listed in boldface indicate discordance. Amyloid-like fragments: bold − strong discordance between T and O including these with biased relation of correlation coefficients − the relation HvO dominating.—cont'd

		RD		Correlation coefficient		
Lyase (1BN8)	**Fragment**	**T-O-R**	**T-O-H**	**HvT**	**TvO**	**HvO**
Other fragments						
β-sheet	168-170*	**0.599**	**0.392**	**−0.161**	**−0.073**	**0.873**
Random coil	1−31	0.469	0.417	0.310	0.642	0.564
Stop − helix	38−46	0.323	0.642	0.773	0.719	0.922
Loop*	63−120	0.640	0.462	0.063	0.252	0.847
Loop*	162−184	0.763	0.682	-0.256	-0.161	0.711
Loop*	203−224	0.684	0.486	-0.093	0.343	0.751
Loop*	297−309	0.690	0.707	0.533	0.739	0.642
RC	327−331	0.035	0.088	0.790	0.993	0.853
RC	354−364	0.416	0.396	0.100	0.657	0.469
Stop − helix	358−364	0.603	0.562	-0.182	0.447	0.548
RC	365−381	0.673	0.748	0.215	0.113	0.761
Helix − parallel	382−393	0.498	0.283	0.390	0.573	0.852
RC	393−399	0.189	0.062	0.433	0.940	0.557
Loops*		0.735	0.610	0.080	0.253	0.802
Helix	2−6	0.510	0.396	0.418	0.352	0.726
	13−17	0.304	0.142	−0.379	0.866	−0.648
	74−79	0.253	0.100	0.801	0.812	0.917
	84−92	0.250	0.184	0.873	0.809	0.922
	93−98	0.675	0.441	−0.198	0.282	0.869
	104−121	0.723	0.551	0.072	0.163	0.846
	207−211	**0.866**	**0.789**	**−0.983**	**−0.742**	**0.835**
	257−261	0.715	0.095	−0.680	0.310	0.276
	385−393	0.352	0.179	0.437	0.679	0.896

* only numbers of residues belonging starting (first) Beta-strand in the Beta sheet are given

upon residues Asp184 and Arg279, along with their immediate neighborhood (5 aa in each direction). Notably, the neighborhood of Asp184 is characterized by a negative value of HvT, indicating that this fragment actively opposes the formation of a hydrophobic core. The larger fragment which contains all catalytic residues (162−280) is likewise discordant and reveals strong involvement of intrinsic hydrophobicity. Given the fragment's role in facilitating catalysis, its status should be regarded as important.

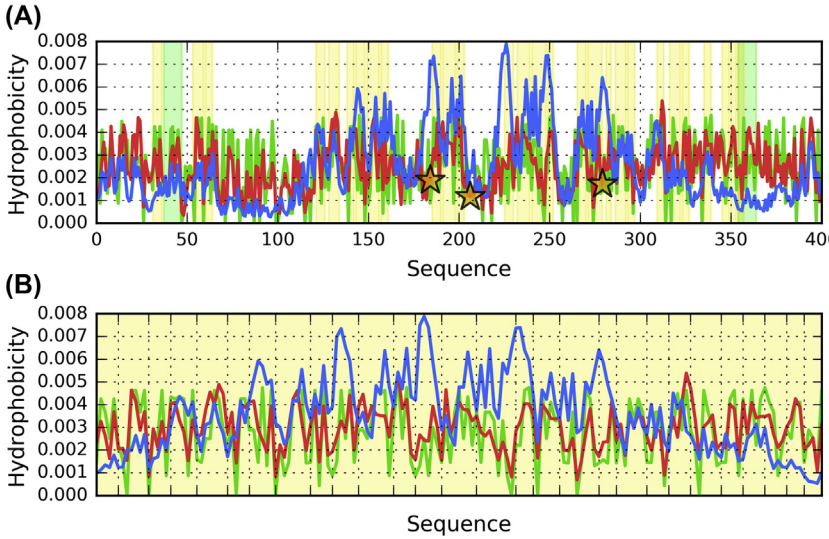


Fig. 7.3 Theoretical (T, blue), observed (O, red) and intrinsic (H, green) hydrophobicity distribution profiles for lyase (1BN8). (A) Complete molecule, (B) Solenoid fragments (β-strands given in Table 7.2, extended by 1 residue in each direction). Background colors highlight the position of solenoid (yellow) and "stop" fragments (green). Orange stars mark positions of catalytic residues. To avoid displaying multiple gaps between the fragments of the solenoid, profiles in B are conflated into contiguous sequence. However ticks on the horizontal axis can be used to determine the location of those fragments withing the complete sequence.

As previously postulated in the Chapter 3, which discusses encoding of information required to ensure biological activity, a highly specific catalytic active site must contain a large quantity of information. This explains the relatively strong discordance between the observed distribution of hydrophobicity and the corresponding deterministic (micelle-like) distribution.

When analyzing the solenoid as a whole, we note that it diverges from the theoretical Gaussian in favor of a distribution dominated by the intrinsic properties of its component residues. In this scope, the β-sheet starting at 122—125 is strongly amyloid-like, with high values of RD along with negative values of both HvT and TvO. Its structure not only diverges from the Gaussian distribution, but actively opposes it.

Fig. 7.4 illustrates the profiles of each fold comprising the 122—125 β-sheet (which is itself part of the solenoid fragment).

The values of FOD parameters describing this β-sheet are consistent with the situation presented in Fig. 7.4A steady decrease in hydrophobicity is predicted by the fuzzy oil drop model (since each successive fold is closer to the

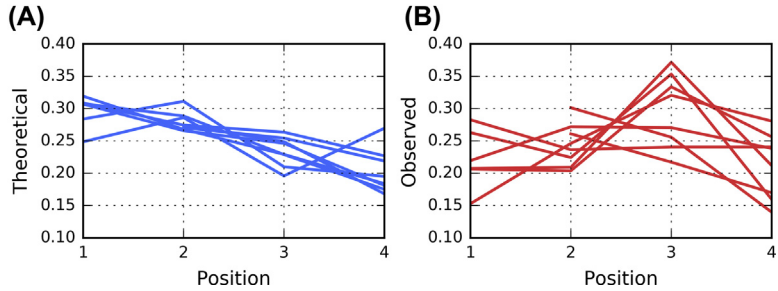


Fig. 7.4 Theoretical (T, blue) and observed (O, red) hydrophobicity distribution profiles for lyase (1BN8) β-sheet which starts with 122—125 strand. (A) Theoretical hydrophobicity, (B) Observed hydrophobicity. Numbers on horizontal axes denote positions of residues relative to sequence fragments of the β-strands after their alignment with their spatial distribution in the β-sheet.

surface of the protein), whereas the observed distribution is characterized by two local maxima separated by a distinct local minimum. The rightmost maximum clearly opposes the theoretical distribution. Thus, Fig. 7.4 reveals the presence of two strongly hydrophobic bands separated by a hydrophilic band. The fact that this band-like pattern is replicated in successive folds, despite differences in their composition, is the hallmark of an amyloid-like structure (this will be further discussed in the next chapter).

It seems that alternating bands of variable hydrophobicity are the driving force behind self-assembly of fibrillar structures. Such structures are, in principle, capable of unrestricted growth, for example by forming complexes with additional lyases. To prevent this undesirable phenomenon, the fragment at 38—46 acts as a "stopper" by presenting a steric obstacle: the helix effectively "caps" the solenoid and enables contact with water thanks to its amphipathic character. The opposite helix at 358—364 lacks these properties, which likely means that it plays a different role, and that a single "cap" is sufficient to prevent unlimited elongation of the solenoid. A similar situation is observed in the previously discussed antifreeze protein.

The band-like pattern exhibited by the 122—125 β-sheet may induce other important effects: for example, it is plausible that it enforces a specific local ordering of water molecules in order to promote catalysis. Similarly, the presence of a Ca^{2+} ion close to the active site may serve to create a suitable force field in terms of electrostatic interactions (the ion, by itself, does not disrupt the distribution of hydrophobicity in its neighborhood). Other fragments — especially random coils — appear to mediate interaction with

the aqueous solvent, given that most of them are found on the surface of the protein and effectively "package" the solenoid.

The status of helical fragments varies, with notable discordances observed in helices which surround the active site (93—98, 104—121, 207—211). Assuming that the molecule as a whole must be able to generate an internal force field which facilitates its biological activity, and that the surrounding solvent must also play a role in this process, these three helices may be regarded as crucially important. Even though our presentation focuses on a single representative lyase, analysis of a larger set of solenoid-containing lyases supports our conclusions. In particular, the properties of the internal force field along with the postulated involvement of the enzyme's aqueous environment (which is acted upon by the protein itself), appear to be of fundamental importance for its biological activity.

Fig. 7.5 visualizes all key fragments of the presented lyase.

In conclusion, it is interesting to note the linear structure of the solenoid fragment, which closely resembles conditions found in amyloids. Considering the threat of unlimited elongation, biologically active solenoids must be equipped with appropriate "caps". The role and structure of such fragments will be further discussed in the Chapter 11. Which presents the potential development of drugs capable of preventing unrestricted growth of amyloid fibrils.

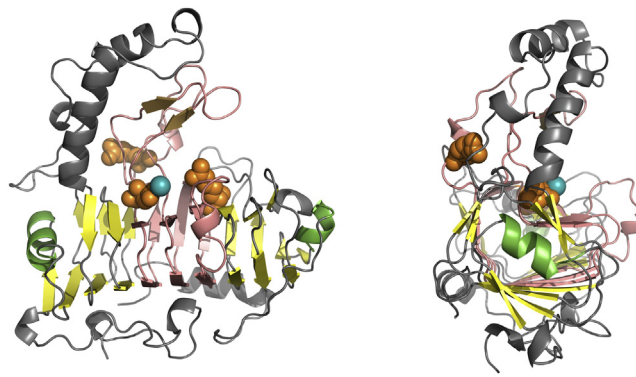


Fig. 7.5 3D presentation of lyase (1BN8) from two angles. Colors distinguish different fragments of the molecule: solenoid — yellow, "stop" fragments — green, neighborhood of catalytic site — salmon (also including parts of the solenoid), orange spheres — catalytic residues, teal sphere — Ca^{2+} ion.

"Perfect" solenoid

A "perfect" solenoid structure is embodied by the antifreeze protein from *Tenebrio monitor*, a beetle species (PDB ID: 1EZG) [6]. The protein consists of 81 residues, including a tandem of 12-residue repeats (TCT_S__C__A_). The regularity of its SS bond structure is striking. FOD analysis reveals that the protein deviates from T in favor of H ($RD > 0.5$) and that therefore its tertiary conformation is determined by the intrinsic hydrophobicity of its residues.

A perfectly symmetrical network of SS bonds (8 in total) can found in the central part of the protein. Consequently, we may speak of a "quasi" hydrophobic core which spans the entire length of the solenoid. The protein also includes several exposed loops consisting of polar residues and likely mediating contact with the solvent (in a manner consistent with the fuzzy oil drop model). Nevertheless, plotting T and O profiles for the entire chain reveals a prominent sinusoidal pattern of alternating maxima and minima, stretching from the N-terminal fragment all the way to the C-terminal fragment.

The distributions plotted in Fig. 7.6 reveal consistent placement of local maxima which appear accordant with the T distribution. The C-terminal fragment is exception representing significant discordance in respect to T distribution.

Another striking property is the near-perfect accordance of fragments bracketed by disulfide bonds. Much like the hydrophobic core, disulfides are thought to contribute to tertiary structural stabilization of proteins. It seems that in the case of 1EZG both factors work in tandem (Figs. 7.7 and 7.8).

The presented protein, which exhibits linear propagation of alternating bands of high and low hydrophobicity, is a good example of how evolution has devised ways to prevent uncontrolled growth of fibrillar structures (which would otherwise be possible in the absence of the N- and C-terminal fragments). The authors of [6] remark that this repetitive sequence translates into an exceptionally regular β-helix, perhaps the most regular protein structure yet observed. Despite the lack of a centric core, the solenoid exhibits strong alignment between the observed and theoretical hydrophobicity, with only its N- and C-terminal fragments remaining locally discordant (Fig. 7.6). The concentration of Cys residues (high intrinsic hydrophobicity) in the central part of the solenoid produces an elongated *quasi*-core, which explains the good agreement between T and O.

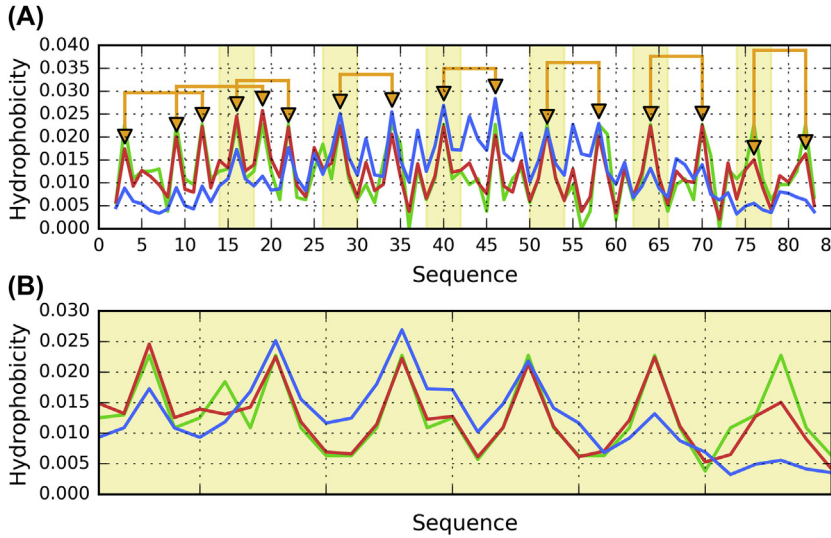


Fig. 7.6 Theoretical (T, blue), observed (O, red) and intrinsic (H, green) hydrophobicity distribution profiles for antifreeze protein (1EZG). (A) Complete molecule, (B) Solenoid fragments (β-strands given in Table 7.3, extended by 1 residue in each direction). Background colors highlight the position of solenoid (yellow). Orange "arrows" correspond to SS bonds formed by Cys residues. To avoid displaying multiple gaps between the fragments of the solenoid, profiles in B are conflated into contiguous sequence. However ticks on the horizontal axis can be used to determine the location of those fragments withing the complete sequence.

From the point of view of the fuzzy oil drop model, this protein is unique in terms of the correlation coefficients computed for successive β-strands and fragments bracketed by disulfide bonds. In all three cases, values of the correlation coefficients are greater than 0.9, indicating near-perfect alignment between the sequence and the structure of the protein. Good agreement between T and O is thought to result from excellent alignment between H and both T and O. Consequently, the protein may be viewed as an example of structural properties encoded in the amino acid sequence. Its antifreeze properties emerge as a result of mechanisms similar to those discussed in Chapter 5: the surface is covered in polar groups which alter the structural properties of the surrounding solvent and disfavor formation of ice crystals. Likewise, the solenoid present in other antifreeze protein (4NU2) exposes alternating bands of high and low hydrophobicity which disrupt the natural structure of the solvent, preventing crystallization.

Table 7.3 FOD parameters (RD values and correlation coefficients) for antifreeze protein (1EZG). Results are listed for the complete molecule, for the β-sheet comprising the solenoid and for fragments bracketed by disulfide bonds. Values listed in boldface indicate discordance.

Antifreeze protein (1EZG)	Fragment	RD		Correlation coefficient		
		T-O-R	T-O-H	HvT	TvO	HvO
Chain	**2—83**	**0.559**	**0.652**	**0.314**	**0.485**	**0.885**
Solenoid						
β-strand	14—18	0.156	0.831	0.945	0.922	0.987
β-strand	26—30	0.183	0.422	0.646	0.916	0.874
β-strand	38—42	0.118	0.879	0.985	0.989	0.995
β-strand	50—54	0.168	0.957	0.994	0.991	0.998
β-strand	62—66	0.202	0.871	0.989	0.994	0.998
β-strand	74—78	0.259	0.650	0.850	0.953	0.895
β-sheet		0.473	0.850	0.508	0.658	0.922
Complete		0.543	0.838	0.332	0.540	0.890
SS-bonds						
	3—12	0.187	0.518	0.866	0.902	0.943
	9—19	0.287	0.788	0.601	0.751	0.950
	16—22	0.338	0.703	0.804	0.723	0.973
	28—34	0.070	0.188	0.831	0.979	0.902
	40—46	0.237	0.416	0.797	0.879	0.950
	52—58	0.344	0.275	0.662	0.779	0.980
	64—70	0.366	0.803	0.546	0.597	0.980
	76—82	0.411	0.624	0.077	0.514	0.855

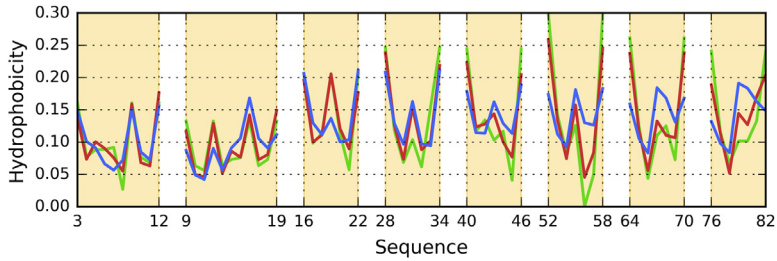


Fig. 7.7 Theoretical (T, blue), observed (O, red) and intrinsic (H, green) hydrophobicity distribution profiles for antifreeze protein (1EZG) fragments bracketed by disulfide bonds. Note that these fragments overlap in the complete sequence, as denoted by Cys positions on horizontal axis.

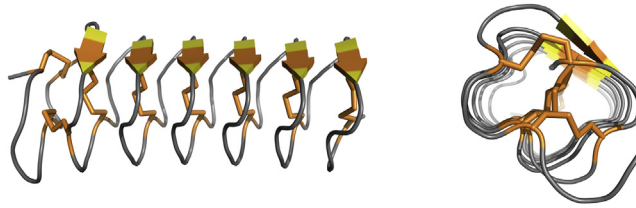


Fig. 7.8 3D presentation of antifreeze protein (1EZG) from two angles. Colors distinguish different fragments of the molecule: solenoid — yellow, disulfides — orange.

Flattened double-walled solenoid

Another solenoid which matches the title of this chapter can be found in a protein listed in PDB under ID 2ZU0: an (E. *coli* sufC-sufD complex involved in the iron-sulfur cluster biosynthesis. The protein is a very large homodimer, with each monomer consisting of two chains. Chain A comprises residues 8—423 (and the same goes for chain B). Chain C has a length of 247 residues and is symmetrical to chain D, although in the latter case only 42 residues are present in PDB file.

Arguably the most interesting component of this protein is the A/B chain dimer, the bulk of which is represented by a large solenoid spanning chain A and continuing in chain B (in a face-on arrangement). The entire complex is characterized by high values of RD and variable correlation coefficients — very low HvT and TvO coupled with very high HvO. This indicates that the listed structure is dominated by intrinsic hydrophobicity (Fig. 7.9).

Overall, the solenoid is structurally similar to amyloids (Fig. 7.10), as evidenced by high values of RD, negative values of both HvT and TvO, and strongly positive values of HvO.

The A/B chain interface is composed of fragments at 355—366 and 340—352 in chain A, along with the corresponding fragments in chain B (antiparallel arrangement). In the context of the solenoid, these fragments are generally accordant with the theoretical distribution, which indicates that hydrophobic interactions play a role in complexation.

Good accordance is also observed for the fragment at 94—105, which functions as a cap. In line with expectations, the outer surface of this fragment faces the environment and mediates entropically advantageous contact with water, preventing unrestricted complexation.

Other fragments exhibit variable deviations from the theoretical distribution, largely resembling an amyloid (the next chapter provides an in-depth

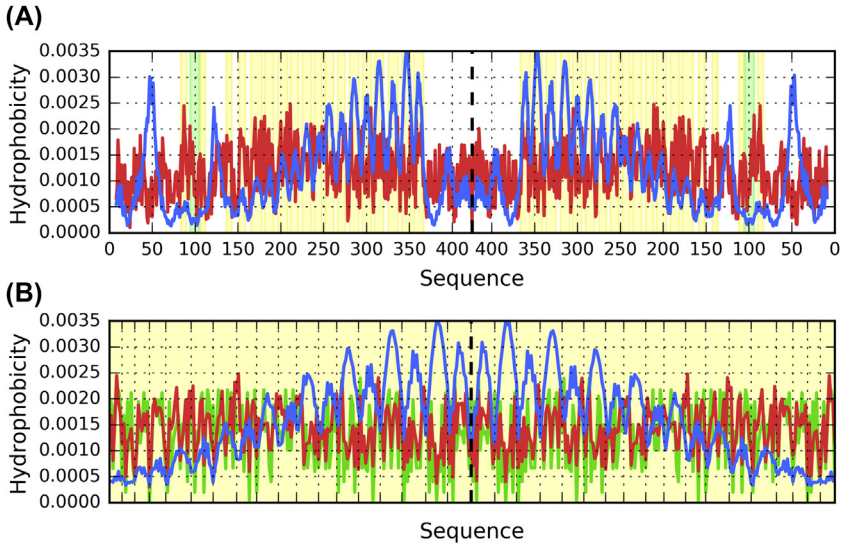


Fig. 7.9 Theoretical (T, blue), observed (O, red) and intrinsic (H, green) hydrophobicity distribution profiles for sufC-sufD complex (2ZU0). (A) Complete molecule, (B) Solenoid fragments (β-strands given in Table 7.4, extended by 1 residue in each direction). Background colors highlight the position of solenoid (yellow) and "stop" fragments (green). Dashed lines separate chain A (left) from chain B (right). To reproduce the symmetry of the complex, results for chain B are present in reverse order. To avoid displaying multiple gaps between the fragments of the solenoid, profiles in B are conflated into contiguous sequence. However ticks on the horizontal axis can be used to determine the location of those fragments withing the complete sequence.

quantitative analysis of amyloid structures in terms of their FOD parameters). In particular, these fragments are dominated by intrinsic hydrophobicity, with high RD coupled with very low values of both TvO and HvT. Such conditions are associated with the propagation of a pattern of alternating bands, as shown in Table 7.4 and Fig. 7.11.

The vast majority of helices observed in chain A (as well as in chain B) remains accordant with the model, suggesting that they play a role in "encapsulating" the solenoid and ensuring that the protein remains soluble. In terms of our model, they act as bridges between the internal (amyloid-like) structure and the external environment, favoring formation of a micelle-like protein, with a predictable hydrophobicity gradient and strongly hydrophilic residues exposed on the surface.

Non-solenoid N-terminal fragment (8—59, 114—134) represents status accordant with idealized distribution while C-terminal fragment

Table 7.4 FOD parameters (RD values and correlation coefficients) for sufC-sufD complex (2ZU0). Results are listed for the complete molecule, for the β-sheets comprising the solenoid, for the solenoid as a whole and for other fragments of the chain, including selected secondary folds. First group of values represents the entire complex as well as each of the four chains analyzed in the context of the complex. "Solenoid (Chain A + Chain B)" corresponds to the status of the solenoid which is formed by chains A and (B).

		RD		Correlation coefficient		
	Fragment	**T-O-R**	**T-O-H**	**HvT**	**TvO**	**HvO**
Complex	A, B, C, D	**0.768**	**0.712**	0.034	0.152	0.754
	Chain A	**0.714**	**0.663**	0.069	0.192	0.762
	Chain B	**0.758**	**0.700**	0.034	0.165	0.754
	Chain C	**0.732**	**0.647**	0.016	0.062	0.793
	Chain D[a]	**0.782**	**0.839**	0.003	-0.126	0.661
Solenoid (Chain A + Chain B)						
β-sheet	**84–89**[b]	**−0.814**	**0.666**	**-0.236**	**-0.112**	**0.830**
β-sheet	**107–111**[b]	**0.831**	**0.781**	**-0.177**	**-0.213**	**0.785**
Complete		**0.804**	**0.703**	**-0.084**	**-0.081**	**0.787**
β-sheet	84−89	0.472	0.735	0.581	0.349	0.818
84-89[b]		0.474	0.772	0.580	0.344	0.841
	152−158	0.686	0.700	-0.438	0.013	0.668
		0.687	0.730	-0.396	-0.019	0.708
	182−191	0.534	0.436	0.154	0.164	0.935
		0.535	0.474	0.156	0.163	0.935
	212−220	0.603	0.242	-0.441	-0.228	0.863
		0.604	0.242	-0.415	-0.207	0.836
	240−247	0.802	0.401	-0.156	-0.401	0.327
		0.791	0.333	-0.181	-0.284	0.121
	267−274	0.507	0.351	0.175	0.250	0.511
		0.528	0.374	0.164	0.172	0.546
	296−305	0.584	0.272	-0.170	-0.284	0.881
		0.587	0.277	-0.175	-0.299	0.888
	326−336	0.615	0.716	-0.088	-0.234	0.894
		0.617	0.697	-0.088	-0.302	0.861
	356−366	0.383	0.215	0.365	0.634	0.929
		0.419	0.247	0.416	0.533	0.972
β-sheet	107−111	0.421	0.305	0.822	0.857	0.878
107-111[b]		0.424	0.253	0.559	0.760	0.785
	137−142	0.585	0.397	0.097	0.345	0.769
		0.650	0.505	0.047	0.102	0.802
	166−177	0.593	0.563	0.158	-0.059	0.686
		0.610	0.565	0.134	-0.114	0.646
	197−207					

(*Continued*)

Table 7.4 FOD parameters (RD values and correlation coefficients) for sufC-sufD complex (2ZU0). Results are listed for the complete molecule, for the β-sheets comprising the solenoid, for the solenoid as a whole and for other fragments of the chain, including selected secondary folds. First group of values represents the entire complex as well as each of the four chains analyzed in the context of the complex. "Solenoid (Chain A + Chain B)" corresponds to the status of the solenoid which is formed by chains A and (B).—cont'd

		RD		Correlation coefficient		
	Fragment	**T-O-R**	**T-O-H**	**HvT**	**TvO**	**HvO**
		0.594	0.539	-0.159	-0.148	0.575
		0.606	0.551	-0.159	-0.172	0.567
	226—235	0.672	0.471	-0.316	-0.382	0.749
		0.686	0.481	-0.327	-0.419	0.756
	252—261	0.635	0.736	-0.631	-0.638	0.839
		0.644	0.734	-0.672	-0.662	0.839
	281—290	0.612	0.526	-0.318	-0.580	0.827
		0.626	0.501	-0.318	-0.625	0.819
	310—319	0.530	0.645	0.216	0.103	0.879
		0.528	0.620	0.215	0.123	0.872
	342—351	0.583	0.333	-0.216	-0.447	0.934
		0.580	0.302	-0.195	-0.368	0.944
Other fragments (CHAIN A)						
Stop − helix + RC	94—105	0.293	0.418	0.427	0.782	0.749
Helices	9—20	0.392	0.291	0.787	0.590	0.565
	26—40	0.419	0.795	0.403	0.528	0.905
	54—59	0.499	0.477	0.632	0.501	0.940
	70—78	0.631	0.577	0.751	0.499	0.774
	95—98	0.334	0.200	0.047	0.711	0.733
	124—134	0.520	0.360	0.040	0.123	0.774
	368—378	0.670	0.665	-0.154	0.113	0.787
	381—399	0.669	0.698	0.317	0.390	0.812
	403—417	0.410	0.227	0.503	0.547	0.938
RC	21—26	0.455	0.527	0.262	0.486	0.917
	39—53	0.424	0.275	0.452	0.466	0.742
	78—90	0.386	0.445	0.538	0.582	0.718
	112—124	0.332	0.255	9,213	0.689	0.739
N-term non-solenoid	8-59 + 114-134	0.394	0.369	0.474	0.769	0.694
C-term non-solenoid	369—423	0.514	0.490	0.300	0.583	0.758

[a]PDB data provides only partial information regarding the structure of chain D.
[b]Numbers correspond to the position of the initial beta fold which belongs to this sheet. Values listed in boldface indicate amyloid-like properties. Underlined values correspond to fragments which "package" the solenoid and are consistent with the fuzzy oil drop model.

3. The solenoid exhibits very strong linear propagation of alternating bands of hydrophobicity, as shown in Fig. 7.11. Negative correlation coefficients indicate that the structure of the presented fragments satisfies amyloid identification criteria. Fig. 7.11 illustrates the peculiar type of discordance exemplified by linear propagation of variably hydrophobic bands in β-sheet starting with strand 107—111 (with short strands omitted). Local minima can be observed in the central part of each strand, where the theoretical model expects hydrophobicity to remain high. Similar distribution is present in the other β-sheet.

(369—423) is discordan (Fig. 7.12) due to a helix (369—377) and a random coil which acts as the P—P interface with chain C and D.

Conclusions and discussion

All of the proteins discussed above exemplify the title of the chapter and each includes fragments characterized by linear propagation of alternating bands of high and low hydrophobicity. While this pattern represents a type of structural ordering, it is unlike the monocentric hydrophobic core expected by the fuzzy oil drop model. When found in biologically active proteins, it calls for a suitable "packaging" which can mediate contact with the aqueous environment (given that the environment promotes a monocentric distribution of hydrophobicity). In principle, linear patterns are capable of unrestricted propagation; however in order to prevent pathologies this capability must be curtailed: thus, solenoid fragments are accompanied by terminal "caps" and — frequently — highly accordant parallel helices, all of which enable entropically advantageous interactions with the solvent (examples may be found by searching PDB for occurrences of "lyase" and "antifreeze" keywords).

In further chapters we will shift our attention to amyloids, which also exhibit band-like linear patterns of hydrophobicity, but lack suitable "packaging" capable of preventing unrestricted growth. These properties will be studied on the example of the beta amyloid (Aβ(1—42) [7,8]) as well as a representative prion amyloid. Similar analysis which focuses on a tau amyloid can be found in Ref. [9].

Clearly, the presence of a solenoid fragment has a profound impact on the structure of the entire protein, and also affects the surrounding aqueous environment. In the case of antifreeze proteins this feedback embodies their biological function, whereas in the presented lyase and the sufC-sufD complex it enables the protein to transmit information to its environment. It is

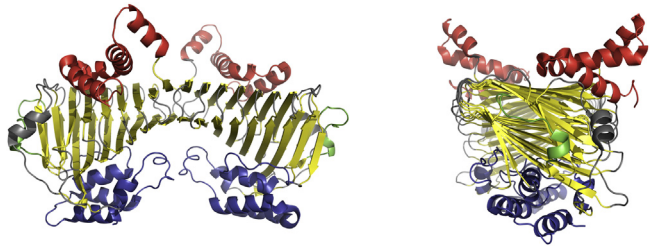


Fig. 7.10 3D presentation of sufC-sufD complex (2ZU0) from two angles. Colors distinguish different fragments of the molecule: solenoid — yellow, "stop" fragments — green, N-terminal non-solenoid structure (8—59, 114—134) — blue, C-terminal non-solenoid structure (369—423) — red.

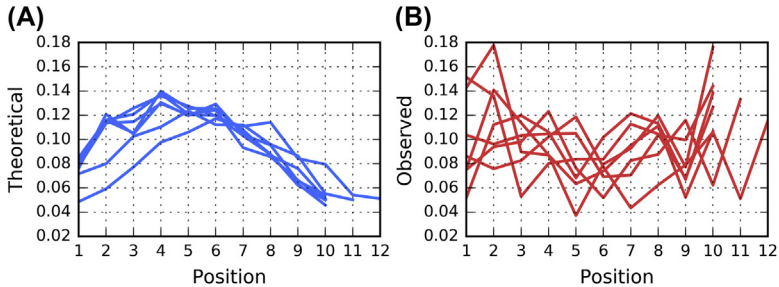


Fig. 7.11 Theoretical (T, blue) and observed (O, red) hydrophobicity distribution profiles for sufC-sufD complex (2ZU0) β-sheet which starts with 107—111 strand. (A) Theoretical hydrophobicity, (B) Observed hydrophobicity.Numbers on horizontal axes denote positions of residues relative to sequence fragments of the β-strands after their alignment with their spatial distribution in the β-sheet.

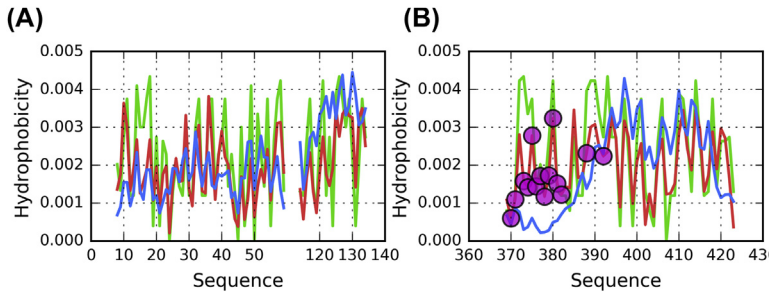


Fig. 7.12 Theoretical (T, blue), observed (O, red) and intrinsic (H, green) hydrophobicity distribution profiles for non-solenoid fragments of chain A of sufC-sufD complex (2ZU0). (A) N-terminal non-solenoid fragment (8—59, 114—134), (B) C-terminal non-solenoid fragment (369—423). Magenta circles mark residues engaged in P—P interaction with chain C.

plausible that the intended ligand is guided to the active site by "sensing" specific local conditions, which are generated by the protein. Such signaling may be regarded as equally important as the influence of the solvent upon the folding process (promoting formation of spherical micelles), while the internal force field — primarily determined by the presence of the solenoid fragment — facilitates catalysis.

The solenoids presented in this chapter may be thought of as ribbonlike micelles. In the Chapter 3. devoted to information theory we suggested that both spherical and ribbonlike micelles carry very little information (which is attributable to their predictable symmetry). Nevertheless, the abovementioned proteins may be regarded as information-rich due to the presence of packaging which encapsulates the solenoid. In contrast, amyloids — devoid of any such packaging — are essentially bare micelles and therefore contain no information. This issue will be further discussed in subsequent chapters.

References

[1] Do H, Kim S-J, Kim HJ, Lee JH. Structure-based characterization and antifreeze properties of a hyperactive ice-binding protein from the Antarctic bacterium Flavobacterium frigorisPS1. Acta Crystallographica Section D Biological Crystallography 2014;70(4): 1061—73. https://doi.org/10.1107/s1399004714000996.

[2] Schutzius TM, Jung S, Maitra T, Graeber G, Köhme M, Poulikakos D. Spontaneous droplet trampolining on rigid superhydrophobic surfaces. Nature 2015;527(7576): 82—5. https://doi.org/10.1038/nature15738.

[3] Macias-Romero C, Nahalka I, Okur HI, Roke S. Optical imaging of surface chemistry and dynamics in confinement. Science 2017;357(6353):784—8. https://doi.org/10.1126/science.aal4346.

[4] Banach M, Konieczny L, Roterman I. Why do antifreeze proteins require a solenoid? Biochimie 2018;144:74—84. https://doi.org/10.1016/j.biochi.2017.10.011.

[5] Pickersgill R, Jenkins J, Harris G, Nasser W, Robert-Baudouy J. The structure of Bacillus subtilis pectate lyase in complex with calcium. Nature Structural & Molecular Biology 1994;1(10):717—23. https://doi.org/10.1038/nsb1094-717.

[6] Liou Y-C, Tocilj A, Davies PL, Jia Z. Mimicry of ice structure by surface hydroxyls and water of a β-helix antifreeze protein. Nature 2000;406(6793):322—4. https://doi.org/10.1038/35018604.

[7] Dułak D, Banach M, Gadzała M, Konieczny L, Roterman I. Structural analysis of the Aβ(15-40) amyloid fibril based on hydrophobicity distribution. Acta Biochimica Polonica 2018. https://doi.org/10.18388/abp.2018_2647.

[8] Banach M, Konieczny L, Roterman I. Fuzzy oil drop model application—from globular proteins to amyloids. Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes 2018:639—58. https://doi.org/10.1007/978-3-319-95843-9_1.

[9] Dułak D, Gadzała M, Banach M, Ptak M, Wiśniowski Z, Konieczny L, Roterman I. Filamentous aggregates of tau proteins fulfil standard amyloid criteria provided by the fuzzy oil drop (FOD) model. International Journal of Molecular Sciences 2018; 19(10):2910. https://doi.org/10.3390/ijms19102910.